Prithvi S. Kandhal,[1] Yiping Wu,[1] Frazier Parker, Jr.,[2] and Peter A. Spellerberg[3]

Precision of Marshall Stability and Flow Test Using 6-in. (152.4-mm) Diameter Specimens

REFERENCE: Kandhal, P. S., Wu, Y., Parker, F., Jr., and Spellerberg, P. A., **"Precision of Marshall Stability and Flow Test Using 6-in. (152.4-mm) Diameter Specimens,"** *Journal of Testing and Evaluation*, JTEVA, Vol. 24, No. 1, January 1996, pp. 20–25.

ABSTRACT: Earlier studies have shown that the repeatability of Marshall stability, flow, and air voids content measurements on 6-in. (152.4-mm) diameter specimens of large stone mixes is better than the repeatability on 4-in. (101.6-mm) diameter specimens. A round robin study involving twelve participating laboratories was conducted to provide information for developing a precision statement for the ASTM Test Method for Resistance to Plastic Flow of Bituminous Mixtures Using Marshall Apparatus (6 inch-Diameter Specimen) (D 5581). Difference two-sigma (d2s) limits were developed to determine acceptable single and multilaboratory differences for bulk specific gravity, percent voids, Marshall stability, and Marshall flow measurements. Analysis of other data collected during the study indicated that stability and flow measurements are not sensitive to minor differences in various 6-in. (152.4-mm) diameter breaking heads currently in use.

KEYWORDS: large stone mixes, Marshall stability, Marshall flow, percent voids, precision statement

The increased incidence of premature rutting of heavy-duty asphalt pavements in recent years is believed to be caused primarily by higher tire pressures and increased wheel loads. Hot mix asphalt (HMA) has served reasonably well in the past, but there is a need to reexamine its design to withstand increased stresses. Most asphalt technologists believe that fundamental changes must be made in the aggregate component of HMA (such as size, shape, texture, and gradation) to reduce rutting to tolerable levels. There is a general agreement that the use of large-size stone in binder and base course will minimize or eliminate rutting of heavy-duty pavements. The term "large stone" is a relative one. For the purpose of this paper large stone is defined as an aggregate with a maximum size of more than 1 in. (25.4 mm).

Marshall mix design procedures are used by 38 out of 50 states according to a survey conducted in 1984 [*1*]. The equipment specified in ASTM Test Method for Resistance to Plastic Flow of Bituminous Mixtures Using Marshall Apparatus (D 1559) consists of a 4-in. (101.6-mm) diameter compaction mold which is intended for mixture containing aggregate up to 1-in. (25.4-mm) maximum size only. The ASTM Test Methods for Resistance to Deformation and Cohesion of Bituminous Mixtures by Means of Hveem Apparatus (D 1560) also use 4-in. (101.6-mm) diameter specimens. This has inhibited the use of HMA containing aggregate larger than 1 in. (25.4 mm) because it cannot be tested with standard mix design equipment.

Other test procedures, such as gyratory compaction, Transport and Road Research Laboratory (TRRL) refusal test, and Minnesota DOT vibrating hammer, have used 6-in. (152.4-mm) diameter molds accommodating $1^1/_2$- to 2-in. (38.1- to 50.8-mm) maximum aggregate size [2]. Of particular interest is the superpave gyratory compactor proposed for Superpave Level 1 design procedure [*3*]. The Strategic Highway Research Program (SHRP) indicated that the Superpave Gyratory Compactor simulates the aggregate particle orientation obtained in the field better than impact-type compactors used in the Marshall procedure. However, it was recently reported [*4*] that the Superpave Gyratory Compactor cannot be used to evaluate dry process rubber asphalt mixtures because of the high resilience of the rubber particles during the compaction process and the time-dependent swelling after compaction of these mixtures. It is expected that Superpave mix design procedures will be modified as they are implemented, and it will be a few years before many state highway agencies fully implement the Superpave design procedures. In the meantime, some highway agencies in the United States and developing countries are likely to continue to use the relatively inexpensive and simple Marshall method.

Literature Review

A study was undertaken by the Pennsylvania Department of Transportation (PennDOT) in the late 1960s to develop equipment and procedures for testing 6-in. (152.4-mm) diameter specimens [*5*]. It was realized that the ASTM Test Method D 1559 (4-in. [101.6-mm] diameter specimen) could not be used for designing Pennsylvania binder course mix and base course mix which specified maximum permissible sizes of $1^1/_2$ and 2 in. (38.1 and 50.8 mm), respectively. Test data indicated that reasonably close compaction levels were achieved in 4- and 6-in. (101.6- and 152.4-mm) diameter molds when the number of blows for the 6-in. (152.4-mm) specimen was 1.5 times that used for the 4-in. (101.6-mm) specimen. Marshall void parameters such as, % voids, % voids in mineral aggregates (VMA), and % voids filled with asphalt (VFA) were also reasonably close. The next step taken by PennDOT in 1970 was to evaluate the repeatability of the test

Manuscript received for publication 4/17/95; accepted for publication 8/7/95.

[1]Associate director and graduate student, respectively, National Center for Asphalt Technology, 211 Ramsey Hall, Auburn University, AL 36849-5354.

[2]Director, Highway Research Center, 233 Harbert Engineering Center, Auburn University, AL 36849-5354.

[3]Manager, AASHTO Materials Reference Laboratory, Bldg. 226, Rm. A365, NIST, Gaithersburg, MD 20899.

results. Statistical analyses indicated better repeatability of 6-in. (152.4-mm) specimens compared with 4-in. (101.6-mm) specimens when testing large stone mixes.

ASTM Subcommittee D04.20 on Mechanical Tests of Bituminous Mixes appointed a task force in December 1988 to develop an ASTM standard test for preparing and testing 6-in. (152.4-mm) diameter Marshall specimens. A draft for this proposed standard was prepared and presented by Kandhal [*6*]. This standard has now been adopted by ASTM as the ASTM Test Method for Resistance to Plastic Flow of Bituminous Mixtures Using Marshall Apparatus (6-inch-Diameter Specimen) (D 5581). This standard follows ASTM Test Method D 1559 which is intended for 4-in. (101.6-mm) diameter specimens except the following significant differences:

1. Equipment for compacting and testing 6-in. (152.4-mm) diameter specimens such as hammer, mold, and breaking head is different. However, the same mechanical compactor unit can be used because the hammer free fall is the same.
2. Since the hammer weighs 22.5 lb (100 N), only a mechanically operated hammer is specified.
3. About 4050 g of mix is required to prepare one 6-in. (152.4-mm) Marshall specimen compared to about 1200 g for a 4-in. (101.6-mm) specimen.
4. The mix is placed in the mold in two approximately equal increments and spading is specified after each increment. Past experience has indicated that this is necessary to avoid honeycombing on the outside surface of the specimen and to obtain the desired density.
5. The number of blows needed for a 6-in. (152.4-mm) diameter and $3^3/_4$-in. (95.25-mm) thick specimen is $1^1/_2$ times the number of blows needed for a 4-in. (101.6-mm) diameter and $2^1/_2$-in. (63.5-mm) thick specimen to obtain equivalent compaction level.

After the preliminary development work done by PennDOT, there was minimal use of 6-in. (152.4-mm) Marshall equipment until 1987. Interest in this equipment was revived when various agencies and producers began testing large stone mixes for minimizing rutting of HMA pavements. According to Kandhal [*6*], several highway agencies and contractors have used 6 in. (152.4 mm) diameter specimens. Kandhal [*6*] obtained comparative test data for 4-in. (101.6-mm) versus 6-in. (152.4-mm) diameter specimens of the same mix from various highway agencies and HMA producers, which indicated that the compaction levels were reasonably close. The data obtained on stability ratio (stability of 6-in. [152.4-mm] specimen/stability of 4-in. [101.6-mm] specimen) and flow ratio (flow of 6-in. [152.4-mm] specimen/flow of 4-in. [101.6-mm] specimen) by various agencies were analyzed. The average stability and flow ratio were determined to be very close to the theoretically derived values of 2.25 and 1.50, respectively. Therefore, it was recommended that the minimum stability requirement for 6-in. (152.4-mm) diameter specimens should be 2.25 times the requirement for 4-in. (101.6-mm) diameter specimens. Similarly, the range of flow values for 6-in. (152.4-mm) specimens should be adjusted to $1^1/_2$ times the values required for 4-in. (101.6-mm) specimens.

A recent study conducted by Kandhal and Brown [*7*] at the National Center for Asphalt Technology, Auburn University, compared 4- and 6-in. (101.6- and 152.4-mm) diameter specimens for testing large stone asphalt mixes. This study compared mix properties such as Marshall stability and flow, indirect tensile strength, and permanent deformation (from static creep test). Both modified Marshall and gyratory testing machines were used to compact the specimens of two sizes. The data indicate larger coefficients of variation when testing 4-in. (101.6-mm) diameter specimens of mix containing aggregate larger than 1 in. (25.4 mm) compared to 6-in. (152.4-mm) specimens. The 6-in. (152.4-mm) diameter specimens also had lower variability in creep test compared to 4-in. (101.6-mm) specimens.

Summary

The literature reviewed in this paper has indicated that compaction levels achieved are reasonably close when the 6- and 4-in. (152.4- and 101.6-mm) diameter compaction molds are used in ASTM Test Methods D 5581 and D 1559 test procedures. Statistical analysis of stability, flow, and air voids showed better repeatability of 6-in. (152.6-mm) specimens used in the ASTM Test Method D 5581 compared to 4-in. (101.6-mm) specimens when testing large stone mixes. However, ASTM Test Method D 5581 did not have the precision statement for repeatability and reproducibility that was needed to be developed from a national round robin testing program. A significant amount of round robin test data is available from Canadian asphalt mix exchange programs [*8*]; however, the available data pertains to 4-in. (101.6-mm) diameter specimens only.

Round Robin Testing Program

A testing program was designed to evaluate the variability of properties measured during the preparation and testing of 6-in. (152.4-mm) diameter Marshall specimens of large stone mix using ASTM Test Method D 5581. Twelve laboratories participated in the study. Materials, specimen preparation procedures, and testing procedures followed during the program are outlined below.

Materials

Samples of aggregate and asphalt cement were distributed by AASHTO Materials Reference Laboratory (AMRL) to the twelve participating laboratories. A viscosity graded AC-20 asphalt cement was used in this round-robin study. A trap rock aggregate from Virginia was used with the gradation in Table 1.

Specimen Preparation

Each participating laboratory was asked to prepare four individual mix specimens, one for each aggregate sample supplied. One

TABLE 1—*Gradation for trap rock aggregate from Virginia.*

Sieve Size	Millimetres	% Passing
2 in.	50	100
$1\,^1/_2$ in.	37.5	100
1 in.	25	77
$^3/_4$ in.	19	65
$^1/_2$ in.	12.5	59
$^3/_8$ in.	9.5	53
#4	4.75	35
#8	2.36	22
#16	1.18	14
#30	0.6	11
#50	0.3	8
#100	0.15	5
#200	0.075	3.3

was used to "butter" the mixing bowl and to provide mix for determining maximum specific gravity. The remaining three batches were used for preparing three replicate Marshall briquets. The three briquets were tested for bulk specific gravity, specimen heights, Marshall stability, and Marshall flow. Each aggregate sample was approximately 4000 g, and each asphalt cement sample was approximately 167 g. When combined by the participating laboratories, these samples produced a mix with 4% asphalt content by weight of mix.

Testing

Laboratories were asked to follow the proposed standard test method (now approved as ASTM Test Method D 5581) for resistance to plastic flow of bituminous mixtures using Marshall apparatus for mixing and compacting 6-in. (152.4-mm) diameter specimens. As noted previously, the number of blows needed for a 6-in. (152.4-mm) diameter specimen is 1.5 times the number of blows needed for a 4-in. (101.6-mm) diameter specimen. Each participating laboratory was asked to use 112 blows (corresponding to 75 blows for a 4-in. (101.6-mm) specimen) to compact the 6-in. (152.4-mm) specimens. A mixing temperature of 300 to 315°F (149 to 156°C) and a compacting temperature of 280 to 290°F (137 to 143°C) were specified.

Data Analysis

Statistical parameters for analyzing round-robin testing data are within- and between-laboratory precision, as described in the ASTM Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials (C 670), the ASTM Practice for Conducting an Interlaboratory Test Program to Determine the Precision of Test Methods for Construction Materials (C 802), and the ASTM Practice for Use of the Terms Precision and Bias in ASTM Test Methods (E 177).

Within-Laboratory Precision

Within-laboratory precision, or repeatability, is a general term for a measure of precision applicable to the variability between test results obtained within a single laboratory by a single operator with a specific set of test apparatus using test specimens taken at random from a single sample of material. However, the "repeatability" as described above is not equivalent to general within-laboratory precision. To obtain an estimate of within-laboratory precision that may reasonably be applied to any laboratory, the whole within-laboratory experiment should be repeated in a number of laboratories. This desired broadly applicable estimate may be obtained by pooling within-laboratory variance from only one operator-day-apparatus combination carried out in each of a number of laboratories. This abbreviated approach, only one operator-day-apparatus combination in each laboratory, is based on the assumption that this estimate of within-laboratory precision does not change, or should not be expected to change, significantly from laboratory to laboratory. Consequently, this measure of precision can be treated as a characteristic of the test method. This pooled or average within-laboratory variance can be given as follows:

$$s_A^2(\text{pooled}) = (\Sigma\, s_i^2)/p \ldots \tag{1}$$

where

s_A^2 = the pooled within-laboratory variance,
s_i^2 = the variance from laboratory i,
p = the total number of laboratories, and
i = a number from 1 to p.

Between-Laboratory Precision

Between-laboratory precision applies to situations in which several laboratories, each with its own operator, apparatus, and environmental conditions, obtain test results on randomly selected specimens from a reasonably uniform sample of material. The variability of the test results may be used to calculate the between-laboratory precision, which, when based on a single test result from each laboratory, is also called the reproducibility of the test method. The between-laboratory component of variance can be calculated as follows:

$$s_{LA}^2 = s_{\bar{x}A}^2 - [s_A^2/n] \ldots \tag{2}$$

where

s_{LA}^2 = between-laboratory component of variance,
$s_{\bar{x}A}^2$ = variance of individual laboratory averages, and
n = the number of replicates tested by each laboratory.

Finally, between-laboratory variance is the sum of pooled within-laboratory variance and the between-laboratory component of variance, and can be calculated as follows:

$$s_{BL}^2 = s_A^2 + s_{LA}^2 \tag{3}$$

Assembling the Data

Table 2 summarizes the data collected from the round robin study. Maximum specific gravity for one sample and the bulk specific gravity, percent voids, Marshall stability, and flow for three samples for all twelve participating laboratories are tabulated.

Processing for Outliners

ASTM Practice C 802, Section 7.5, discourages discarding individual test results that appear to differ by suspiciously large amounts from the others, unless there is a clear evidence of some physical reason to consider the result faulty. For this study, a three standard deviation ($\pm 3s$) criterion was used for evaluating outliers. Means ($\bar{x}$) and standard deviations (s) were computed for each parameter (n = 36). The maximum absolute differences $|x - \bar{x}|$ were computed for each parameter and compared to $3s$.

Bulk Specific Gravity—$s = 0.020$, $3s = 0.060$, $\bar{x} = 2.662$, and maximum $|x - \bar{x}| = 2.711 - 2.662 = 0.049 < 0.060$.

% Voids—$s = 0.76$, $3s = 2.28$, $\bar{x} = 3.90$, and maximum $|x - \bar{x}| = 3.90 - 2.21 = 1.69 < 2.28$.

Marshall Stability—$s = 892$ lb (3968 N), $3s = 2676$ lb, $\bar{x} = 5878$ lb (26 147 N), and maximum $|x - \bar{x}| = 7616 - 5878 = 1739$ lb < 2676 lb.

Flow—$s = 4.3$ (0.01 in.) or 1.09 (mm), $3s = 12.9$ (0.01 in.), $\bar{x} = 18.6$ (0.01 in.) or 4.72 (mm), and maximum $|x - \bar{x}| = 29 - 18.6 = 10.4 < 12.9$.

TABLE 2—*Summary of round robin data.*

Lab.	Max. Sp. Gr.	Bulk Sp. Gr.	% Voids	Stability, lb (N)	Flow, 0.01 in. or 0.25 mm
01	2.770	2.655	4.16	5712 (25 408)	13.5
		2.676	3.40	7616 (33 877)	20.0
		2.677	3.37	5186 (23 068)	21.0
02	2.776	2.649	4.57	7370 (32 783)	18.0
		2.651	4.50	5018 (22 321)	18.0
		2.662	4.11	6966 (30 986)	21.0
03	2.781	2.665	4.17	7245 (32 227)	29.0
		2.662	4.26	7130 (31 716)	24.0
		2.657	4.46	6210 (27 623)	23.0
04	2.763	2.639	4.49	5800 (25 800)	14.0
		2.666	3.51	5400 (24 020)	16.0
		2.645	4.27	5180 (23 042)	13.0
05	2.760	2.654	3.84	5250 (23 353)	17.0
		2.649	4.02	5350 (23 798)	18.5
		2.662	3.55	4850 (21 574)	20.5
06	2.777	2.685	3.31	6550 (29 136)	24.0
		2.711	2.38	5650 (25 132)	22.0
		2.691	3.10	5175 (23 019)	28.0
07	2.760	2.672	3.20	5700 (25 355)	18.0
		2.670	3.20	7100 (31 582)	28.0
		2.649	4.00	5700 (25 355)	18.0
08	2.763	2.697	2.39	6540 (29 091)	14.0
		2.696	2.42	7392 (32 881)	19.0
		2.702	2.21	6160 (27 401)	12.0
09	2.772	2.641	4.73	5432 (24 163)	17.0
		2.640	4.76	5040 (22 419)	14.0
		2.640	4.75	6270 (27 890)	21.0
10	2.766	2.636	4.70	4860 (21 618)	16.0
		2.633	4.80	4220 (18 771)	14.0
		2.634	4.80	4890 (21 752)	16.0
11	2.789	2.656	4.77	6950 (30 915)	19.1
		2.663	4.50	5960 (26 511)	15.9
		2.662	4.54	6150 (27 356)	17.6
12	2.769	2.667	3.68	5090 (22 641)	14.2
		2.671	3.54	4820 (21 440)	14.3
		2.664	3.79	5690 (25 310)	19.5

TABLE 3—*Statistical parameters for each laboratory.*

	Bulk Specific Gravity		% Voids in Total Mix		Marshall Stability		Marshall Flow	
Lab.	$\bar{x}_i$	s_i^2	$\bar{x}_i$	s_i^2	$\bar{x}_i$	s_i^2	$\bar{x}_i$	s_i^2
01	2.669	0.000 154	3.64	0.2004	6171	1 634 465	18.2	16.58
02	2.654	0.000 049	4.39	0.0614	6451	1 581 637	19.0	3.00
03	2.661	0.000 016	4.30	0.0220	6862	321 808	25.3	10.33
04	2.650	0.000 201	4.09	0.2644	5460	98 800	14.3	2.33
05	2.655	0.000 043	3.80	0.0562	5150	70 000	18.7	3.08
06	2.696	0.000 185	2.93	0.2379	5792	487 708	24.7	9.33
07	2.664	0.000 162	3.47	0.2133	6167	653 333	21.3	33.33
08	2.698	0.000 010	2.34	0.0129	6697	398 021	15.0	13.00
09	2.640	0.000 000	4.75	0.0002	5581	369 181	17.3	12.33
10	2.634	0.000 002	4.77	0.0033	4657	143 233	15.3	1.33
11	2.660	0.000 014	4.60	0.0212	6353	276 033	17.5	2.56
12	2.667	0.000 012	3.67	0.0157	5200	198 300	16.0	9.19

Each laboratory performed only one maximum mix specific gravity test. These data were also analyzed as follows for outliers (n = 12): s = 0.009, $3s$ = 0.027, $\bar{x}$ = 2.771, and maximum $|x - \bar{x}|$ = 2.789 − 2.771 = 0.018 < 0.027.

Based on the above calculations with (± $3s$) criteria, there were no outliers for the data in this study.

Computation of Statistical Parameters

Statistical parameters for each laboratory are summarized in Table 3. Mean ($\bar{x}_i$) and variance (s_i^2) for each laboratory (n = 3) are computed using standard statistical methods. Using this data and procedures described previously from ASTM Practice C 802, within-laboratory variance (s_A^2) and between-laboratory variance (s_{BL}^2) were computed for each mix property and are shown in Table 4. Also shown in Table 4 are means, standard deviations, and coefficients of variation.

Investigation of Homogeneity of Variance

ASTM Practice C 802 contains a procedure for determining if variances for the different laboratories in a study are the same (homogeneity of variance). High values of variance are examined by comparing the ratio of the largest variance (for an individual laboratory) to the sum of the variances for all laboratories. There is agreement among variances if the computed ratio is less than criteria set at a 5% significance level. For n = 12 laboratories and p = 3 replicates, the limiting ratio is 0.3924.

Small variances tend to be not as troublesome as large variances. However, should a laboratory conduct the testing in such a manner that normal variation does not occur, the variation will be unrealistically low. If no significantly high variance is present, ASTM Practice C 802 contains criteria for judging the significance of the lowest variation. The statistic used is the ratio of the highest to lowest variance. There is agreement among variances when this ratio is less than 704 (n = 12, p = 3, and α = 5%).

Ratios of the largest variance to the sum of the variance are as follows:

bulk specific gravity: 0.2363,
% voids: 0.2384,
Marshall stability: 0.2611, and
Marshall flow: 0.2863.

Since those ratios are less than 0.3924, it was concluded that the largest variance was not significantly different from the others.

Ratios of the largest to the smallest variance are as follows:

bulk specific gravity: 670,
% voids: 1322,
Marshall stability: 23, and
Marshall flow: 26.

The above ratios indicate that the variances for Laboratory 9 may be unrealistically low for % voids.

TABLE 4—*Summary of statistical parameters for total study.*

	Mix Properties			
	Bulk Specific Gravity	% Voids	Stability, lb (N)	Flow, 0.01 in. or 0.25 mm
$\bar{x}$	2.662	3.90	5 878 (26 147)	18.6
s_A^2	0.000 071	0.0924	521 627 (10 321 164)	9.69
s_{BL}^2	0.000 416	0.6157	812 585 (16 078 199)	19.36
s_A	0.008	0.30	722 (3217)	3.1
s_{BL}	0.020	0.78	901 (4010)	4.4
$s_A/\bar{x}$* 100	0.3	7.7	12.3	16.7
$s_{BL}/\bar{x}$* 100	0.75	20.0	15.3	23.7

Development of Precision Statements

A precision statement meeting the requirements of ASTM Practice C 670 normally contains two main elements described as follows.

1. *Single-Operator Precision*—A measurement of the greatest difference between two results that would be considered acceptable when properly conducted repetitive determinations are made on identical materials by a competent operator.
2. *Multilaboratory-Precision*—A measurement of the greatest difference between two test results that would be considered acceptable when properly conducted determinations are made by two different operators in different laboratories on portions of a material that are intended to be identical, or nearly identical as possible.

As described in ASTM Practice C 670, precision statements are based on within-laboratory and between-laboratory standard deviations. Single-operator requirements are based on within-laboratory standard deviations and multilaboratory requirements are based on between-laboratory standard deviations. Limits for acceptable differences between two test results are obtained by multiplying the appropriate standard deviation by $2\sqrt{2}$ to obtain the difference two-sigma limit (d2s). The d2s limit is the difference between two test results that would be equal or exceed in only one case in twenty for normal testing.

The standard deviations in Table 4 for bulk specific gravity, % air voids, Marshall stability, and Marshall flow are retabulated and used to compute the difference two-sigma limits (d2s) in Table 5. Using Marshall stability as an example, Table 5 can be interpreted as follows:

The single-operator standard deviation is 722 lbs (3212 N) for measured Marshall stabilities ranging from 4657 lbs (20 715 N) to 6862 lbs (30 524 N). Therefore, the results of two properly conducted tests by the same operator on the same mix should not differ by more than the difference two-sigma limit of 2042 lbs (9083 N).

The multilaboratory standard deviation is 901 lbs (4008 N). Therefore, the results of properly conducted tests from two different laboratories on the same mix should not differ by more than 2548 lbs (11 334 N).

TABLE 5—*Criteria for judging acceptability of bulk specific gravity, % voids, Marshall stability, and flow measurements on 6-in. (25.4-mm) diameter specimens.*

Parameters	Standard Deviations (1s)	Difference Two-Sigma Limit (d2s)
Bulk specific gravity		
single operator	0.0084	0.0237
multilaboratory	0.0204	0.058
% Air voids		
single operator	0.30	0.86
multilaboratory	0.78	2.21
Stability (lb-force)		
single operator	722	2042
multilaboratory	901	2548
Flow (0.01-in.)		
single operator	3.1	8.8
multilaboratory	4.4	12.4

The precision statement can also be expressed in terms of the coefficient of variation rather than absolute values of Marshall stability and flow, which are likely to be different for different asphalt mixtures. Table 6 shows the precision statement for stability and flow based on the coefficient of variation.

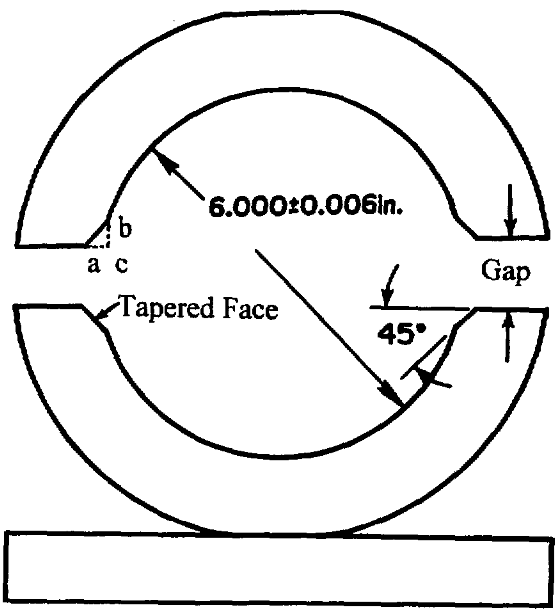
Sensitivity of Stability and Flow to Breaking Head Gap and Bevel Width

Different equipment manufacturers have used slightly different gaps and bevel dimensions for their 6-in. (152.4-mm) diameter breaking heads before the test was standardized as shown in Fig. 1. Therefore, a secondary problem to be investigated in this study were the effects of gap and bevel (tapered face) width of the breaking head on the stability and flow. Kandhal [1] reported that stability for 4-in. (101.6-mm) diameter specimens may be sensitive

TABLE 6—*Acceptability of test results for stability and flow.*

Test and Type Index	Coefficient of Variation (% of mean)[a]	Acceptable Range of Two Test Results (% of mean)[a]
Single-operator precision:		
Stability, pounds-force	12.3	34.8
Flow, hundredth of an inch	16.7	47.2
Multilaboratory precision:		
Stability, pounds-force	15.3	43.3
Flow, hundredth of an inch	23.7	67.0

[a]These numbers represent, respectively, the (1S%) and (D2S%) limits as described in ASTM Practice C 670.



Combined Gap = Gap + 2 bc = Gap + 2 ab / $\sqrt{2}$
= Gap + $\sqrt{2}$ ab
Where ab = Width of Tapered Face

FIG. 1—*Geometry of breaking head. 1 in. = 25.4 mm.*

to breaking head geometry. A questionnaire sheet was sent to each participating laboratory requesting relevant information. Only six laboratories provided complete information on both gap and bevel width for the breaking head used in the testing. The combined gap varied from 1.17 to 1.61 in. (29.72 to 40.89 mm). Since the stability and flow are believed to be affected by the extent of confinement, the combined gap and bevel width were used in the analysis of sensitivity of stability and flow. Data are plotted in Figs. 2 and 3 for stability and flow, respectively. These plots show no apparent relationships between stability or flow and combined gap and bevel width. This implies that stability and flow are not sensitive to the different 6-in. (152.4-mm) diameter breaking heads currently on the market.

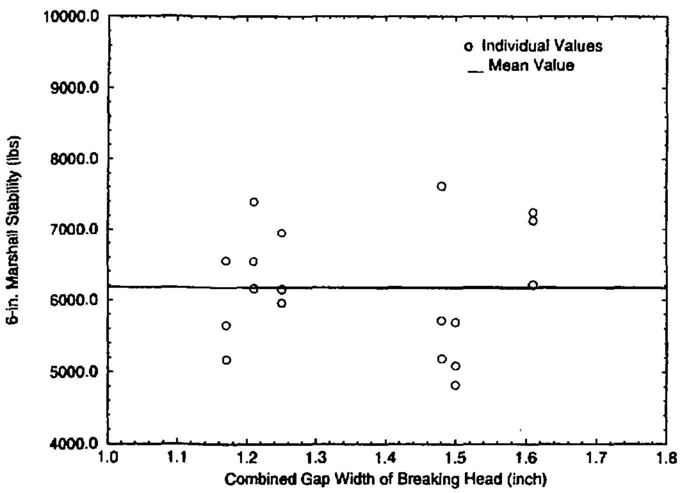


FIG. 2—*Combined gap width of breaking head versus Marshall stability. 1 in. = 25.4 mm and 1 lbf = 4.448 N.*

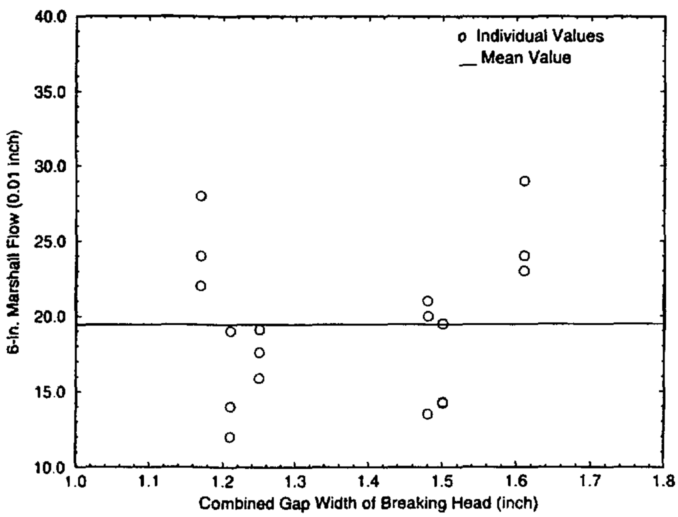


FIG. 3—*Combined gap width of breaking head versus Marshall flow. 1 in. = 25.4 mm.*

Conclusions

The literature review indicated that the 6-in. (152.4-mm) diameter specimen Marshall apparatus has less variability in the test results than the 4 in. (101.6 mm) and therefore should be used for the design of large stone asphalt mixes. A round robin study was conducted with the cooperation of AASHTO Materials Reference Laboratory to provide information for developing precision statements when testing 6-in. (152.4-mm) diameter Marshall specimens using the new ASTM Test Method D 5581. Difference two-sigma (d2s) limits were developed to determine acceptable single-operator and multilaboratory differences for bulk specific gravity, % air voids, Marshall stability, and Marshall flow. Analysis of data from the round robin study also indicated that the stability and flow values are not sensitive to the minor differences in the geometry of 6-in. (152.4-mm) diameter breaking heads currently in use.

References

[1] Kandhal, P. S. and Koehler, W. S., "Marshall Mix Design Method: Current Practices," *Proceedings, Association of Asphalt Paving Technologists*, Vol. 54, 1985.

[2] Acott, M., "The Design of Hot Mix Asphalt for Heavy Duty Pavements," QIS 111/86, *National Asphalt Pavement Association*, Oct. 1987.

[3] "Superior Performance Pavement-SUPERPAVE™, The Product of the SHRP Asphalt Research Program," Strategic Highway Research Program, National Research Council, March 1993.

[4] Hafez, I. H. and Witczak, "Comparison of Marshall and SUPERPAVE™ Level I Mix Design for Asphalt Mixes," paper prepared for presentation at 74th Annual Meeting of Transportation Research Board, 1995.

[5] "Comparison of 4 and 6-in. Diameter Molded Specimens," Pennsylvania Department of Transportation. Bureau of Materials, Testing and Research, Status Report, 21 Feb. 1969.

[6] Kandhal, P. S., "Large Stone Asphalt Mixes: Design and Construction," *Journal of the Association of Asphalt Paving Technologists*, Vol. 59, 1990.

[7] Kandhal, P. S. and Brown, E. R., "Comparative Evaluation of 4-Inch and 6-Inch Diameter Specimens for Testing Large Stone Asphalt Mixes," in *Proceedings of the First Materials Engineering Congress*, Vol. 1, American Society of Civil Engineers, 1990.

[8] Beaudoin, M., Langlois, P., and Proteau, D., "Programme d'Echange Canadien 1991–1992 sur les Enrobés Bitumineux," *Canadian Technical Asphalt Association Proceedings*, Vol. XXXVII, 1992.